

Survey Methodology FY 2001 - Introduction

from: E.D. TAB: Public Libraries in the United States: Fiscal Year 2001

The tables in this report summarize information about public libraries in the 50 states and the District of Columbia for state fiscal year (FY) 2001. Forty-nine states, the District of Columbia, and two outlying areas (Guam and the U.S. Virgin Islands) submitted data for FY 2001.¹⁵ Data from Guam and the U.S. Virgin Islands are included in the tables, but not in the table totals. Minnesota did not respond to the survey—all of its data are imputed.¹⁶ The data were collected through the Public Libraries Survey (PLS), conducted annually by the National Center for Education Statistics (NCES) through the Federal-State Cooperative System (FSCS) for Public Library Data. The FY 2001 survey is the 14th in the series.¹⁷ This report is based on the final data file.

This report includes information about service measures such as access to the Internet and other electronic services, number of Internet terminals used by staff only, number of Internet terminals used by the general public, reference transactions, public service hours, interlibrary loans, circulation, library visits, children's program attendance, and circulation of children's materials. It also includes information about size of collection, staffing, operating income and expenditures, type of geographic service area, type of legal basis, type of administrative structure, and number and type of public library service outlets. Data were imputed for nonresponding libraries. More detail on the methodology used for imputation is provided in the Technical Notes section of this report. Rankings of states on selected characteristics are included in appendix A.

FSCS is a cooperative system through which states and the outlying areas submit individual public library data to NCES on a voluntary basis. At the state level, FSCS is administered by State Data Coordinators (SDCs) appointed by the Chief Officers of State Library Agencies (COSLA). The SDC collects the requested data from public libraries and submits these data to NCES. NCES aggregates the data to provide the state and national totals presented in this report.

There are 61 tables in this report, as follows: Table 1 has 3 parts (tables 1, 1A, and 1B); tables 2 through 24 are in sets of 2 each; and appendix A includes 10 tables of state rankings by selected data items. The base table in each set (tables 2 through 24) displays data for the nation as a whole, and for each of the 50 states, the District of Columbia, and 2 outlying areas. The "A" table in each set displays the same data by 11 ranges of population of legal service area.

Items Collected but Omitted from This Report

Newer items. Number of Users of Electronic Resources Per Typical Week was collected beginning with FY 99. This item and the items listed below are not included in this report but are on the data file.

- Information on public library service outlets, including the location of public library service outlets relative to a metropolitan area, number of books-by-mail only outlets, number of bookmobiles by bookmobile outlet, and Web address.
- Identifying information about individual public libraries and their outlets, such as address and telephone number.
- The public library's status as an FSCS public library. (The survey includes the question, "*Does this public library meet all the criteria of the FSCS public library definition?*")

¹⁵Data were not reported by the following outlying areas: American Samoa, the Northern Marianas, Palau, and Puerto Rico.

¹⁶The imputation (estimation) of missing data is discussed in appendix C.

¹⁷Trend data from some of the earlier surveys are discussed in *Public Library Trends Analysis, 1992–1996* (NCES 2001–324) (Glover, 2001), a Statistical Analysis Report released by NCES in the summer of 2001.

Terminology Used to Describe the Structures and Relationships Among and Within Public Libraries

In this report, the term public library means an administrative entity. See appendix D for the definition of public library.

- **Administrative entities.** An administrative entity is the agency that is legally established under local or state law to provide public library service to the population of a local jurisdiction. The administrative entity may have a single public library service outlet, or it may have more than one public library service outlet. The types of administrative structures for public libraries are reported in table 23.
- **Public library service outlets.** Public libraries can have one or more outlets that provide direct service to the public. The three types of public library service outlets included in this report are central library outlets, branch library outlets, and bookmobile outlets. Information on a fourth type of outlet, books-by-mail only outlets, was collected but omitted from this report. See appendix D for definitions of these terms. Table 2 reports data concerning public library service outlets.

Caveats for Using These Data

Using the Data to Make Comparisons

The data include imputations, at the unit and item levels, for nonresponding libraries. Comparisons to data prior to FY 92 should be made with caution, as earlier data do not include imputations for nonresponse, and the percentage of libraries responding to a given item varied widely among states.

It is important to make state data comparisons with caution because of differences in reporting periods (see following table) and adherence to survey definitions.¹⁸ The District of Columbia, while not a state, is included in this report. Special care should be used in comparing the District's data to state data since it is an urban area, not a state. Caution should also be used in making comparisons with the state of Hawaii, as Hawaii reports only one public library for the entire state. Minnesota did not respond to the survey—all of its data are imputed. See appendix C for a discussion of the imputation methodology for nonresponding libraries.

Reporting Period

The FY 2001 PLS requested data for state fiscal year 2001. In some states, the fiscal year reporting period varies among localities. These states were requested to report the earliest starting date and latest ending date reported by their public libraries. Although the reporting period for some states spans more than a 12-month period (see table below) due to multiple fiscal cycles, each public library provided data for a 12-month period.

States by Reporting Period

07/00 to 06/01		01/01 to 12/01	Other
AK	MO	AR	12/99 to 09/01: MI
AZ	MT	CO	01/00 to 06/01: VT
CA	NC	IN	03/00 to 12/01: NY
CT	NM	KS	06/00 to 12/01: PA
DE	NV	LA	07/00 to 12/01: NH, UT
GA	OK	ME	10/00 to 09/01: AL, DC, FL, ID, MS, GU*, VI*
HI	OR	MN	12/00 to 12/01: NE
IA	RI	ND	02/00 to 12/01: TX
IL	SC	NJ	
KY	TN	OH	
MA	VA	SD	
MD	WV	WA	
	WY	WI	

*GU—Guam;
VI—Virgin Islands

¹⁸The definitions used by some states in collecting data from their public libraries may not be consistent with the PLS definitions. The 1994 NCES *Report on Coverage Evaluation in the Public Library Statistics Program* (NCES 94-430) and the 1995 NCES *Report on Evaluation of Definitions Used in the Public Library Statistics Program* (NCES 95-430) address issues of consistency in definitions among states. The reports are available on the NCES web site at <http://nces.ed.gov/pubsearch/getpubcats.asp?sid=041#052>

Definitions

The survey definitions are included in appendix D. A few key items are discussed below.

Public library. The definition of public library in the Public Libraries Survey (PLS) is as follows: “A public library is an entity that is established under state enabling laws or regulations to serve a community, district, or region, and that provides at least the following: (1) an organized collection of printed or other library materials, or a combination thereof; (2) paid staff; (3) an established schedule in which services of the staff are available to the public; (4) the facilities necessary to support such a collection, staff, and schedule; and (5) is supported in whole or in part with public funds.” The survey data file identifies the status of each public library based on the survey definition (97 percent of the public libraries meet the definition). The remaining libraries are included on the data file because they meet the definition of a public library under state laws, which may vary. It is not known which parts of the definition of public library in the PLS are not met by these remaining libraries.

Library visits and reference transactions. Public libraries reported annual library visits and annual reference transactions based on actual counts, if available. Otherwise, annual estimates were provided based on a typical week in October, multiplied by 52. See the definitions of these items in appendix D for more information.

Population of legal service area. The PLS has three population items: (1) Population of Legal Service Area for each public library, (2) Total Unduplicated Population of Legal Service Areas in the state that receive public library services, and (3) Official State Total Population Estimate. The data for all population items are provided by the state library agency. There are significant methodological differences in the ways states calculate the first two items, and the time periods these counts were made for also vary among states.

The total Population of Legal Service Area for all public libraries in a state may exceed the state's Total Unduplicated Population of Legal Service Areas or the Official State Total Population Estimate. This occurs when the state has one or more geographically adjacent libraries (for example, a county library and a city library within the county) that serve, and therefore count, the same population. Twenty-six states had such overlapping service areas in FY 2001. (See appendix B for a list of these states.)

In order to make meaningful state comparisons using total Population of Legal Service Area data (for example, the number of books/serial volumes per capita), the data were adjusted to eliminate duplicative reporting due to overlapping service areas. The public library data file has a derived unduplicated population of legal service area for *each library* for this purpose (the variable is called POPU_UND). This value was prorated for each library by calculating the ratio of a library's Population of Legal Service Area to the state's total Population of Legal Service Area and applying the ratio to the state's Total Unduplicated Population of Legal Service Areas. Table 1 provides the total Population of Legal Service Area and the Total Unduplicated Population of Legal Service Areas for each state.¹⁹

¹⁹The unduplicated population of legal service areas provided by the states does not include unserved areas and may vary from data provided by sources using standard methodology (e.g., U.S. Census Bureau).

Appendix C—Technical Notes

Survey Universe

The survey frame includes 9,133 public libraries (9,129 public libraries in the 50 states and the District of Columbia and four public libraries in the outlying areas of Guam, the Northern Mariana Islands, Palau, and the U.S. Virgin Islands). The public libraries were identified by the state library agencies.¹ Although data were not systematically collected from public libraries on Native American reservations, a total of 35 such libraries were reported in fiscal year (FY) 2001. The survey population does not include military libraries that provide public library services or libraries that serve the residents of institutions. The FY 2001 survey is the 14th in the series. This report is based on the final data file.

Survey Response

Unit response. A total of 8,808 of the 9,133 public libraries in the survey frame responded to the FY 2001 PLS (8,806 public libraries in the 50 states and the District of Columbia and two in the outlying areas), for a unit response rate of 96.4 percent. Minnesota did not respond to the survey—all of its data are imputed (estimated). Missing data were not imputed for nonresponding outlying areas. Respondents to the survey are defined as public libraries for which the following data were reported: population of the legal service area and at least three of the five following items: total paid employees, total operating income, total operating expenditures, book/serial volumes, and total circulation.

Total response. The base for calculating response rates for specific items is the total number of libraries in the survey frame, including unit nonrespondents. For national totals, total response rates did not fall below 85 percent for any items in this report.² For state totals, total response rates fell below 85 percent in 24 states for one or more items. (Note: The total response rates are included in the tables in this report.) Missing data were imputed and included in the state and national totals. See *Imputation* below for a discussion of the imputation methodology.

Calculations Reported in Tables

- Percentages rather than raw numbers are used in many of the tables to provide a clearer picture of patterns in the data. Percentage distributions may not sum to 100 due to rounding. To obtain a raw number from a percentage distribution table, multiply the percentage for the item by the total for the item (the total may be on a different table). For example, in table 9, the number of public libraries with less than 5,000 book and serial volumes is 338 (9,129 x .037).
- Selected tables include *per capita* values for some items and *per 1,000 population* or *per 5,000 population* values for others (e.g., table 4). The value that yielded the most straightforward presentation of the data was used.

Paid Staff (Full-Time Equivalent)

Paid staff were reported in full-time equivalents (FTEs) (tables 10 through 11A). To ensure comparable data, 40 hours was set as the measure of full-time employment (for example, 60 hours per week of part-time work by employees in a staff category divided by the 40-hour measure equals 1.50 FTEs). FTE data were reported to two decimal places (rounded to one decimal place in the tables).

¹Public libraries in two outlying areas, American Samoa and Puerto Rico, are not included in the survey frame because their state library agencies have never responded to NCES's request for participation in the survey. Their public libraries have not been identified and cannot be included in the response rate calculations.

²The NCES Statistical Standards stipulate that if the item response rate is below 85 percent for any items used in a report, a nonresponse bias analysis is required for each of those items.

Data Collection and Use of Technology

The FY 2001 PLS was released to the states over the Internet on November 15, 2001. States were placed into one of three reporting groups (with survey due dates of April 15, July 31, or August 31, 2002), based on their fiscal cycles and whether or not they claimed an extraordinary reporting hardship. States reported their data using a personal computer data collection software program called WinPLUS (Windows Public Library Universe System) which they downloaded from the Internet. WinPLUS was developed for NCES by the U.S. Census Bureau (the data collection agent). Edit follow-up was completed in mid-December of 2002. The editing process is described below.

Editing

State level. The survey software has an edit check program that generates on-screen warnings during the data entry/import process, enabling the respondent to review the data and correct many errors immediately. Following data entry/import, the respondent can generate an on-screen or printed edit report for further review and correction of the data before submitting the final file to NCES. Four types of edit checks were performed:

1. *Relational edit checks.* This is a data consistency check between related data elements. For example, an edit check message is generated if the number of ALA-MLS Librarians is greater than Total Librarians.
2. *Out-of-range edit checks.* This is a comparison of data reported for an item to the "acceptable range" of numeric values. For example, an edit check message is generated if average Public Service Hours per outlet per week is less than 10, or if Total Circulation reported this year is not within $\pm 5,000$ or $\pm 10\%$ to $\pm 25\%$ of last year's value for Total Circulation.
3. *Arithmetic edit checks.* This is an arithmetic check comparing a reported total to the sum of its parts. For example, an edit check message is generated if Total Operating Income is not equal to the sum of its parts (Local Government Income, State Government Income, Federal Government Income, and Other Income).
4. *Blank, zero, or invalid data edit checks.* This is a check of reported data against acceptable values. For example, an edit check message is generated if Book/Serial Volumes is 0 or blank.

Respondents also used the survey software to generate state summary tables (corresponding to the tables in this report but limited to their state's data) and single-library tables (showing data for individual public libraries in their state). States were encouraged to review the tables for data quality before submitting their final data to NCES. States submitted their final data with a signed form from the Chief Officer of the State Library Agency certifying the accuracy of the data.

National level. The U.S. Census Bureau and NCES reviewed and edited the state data submissions, working closely with the state data coordinators and the FSCS Steering Committee.

Imputation

All libraries, including nonresponding libraries, were sorted into imputation cells based on the region and size of population served. Item imputation was performed on each record with nonresponse variables. The data are identified as either imputed (estimated) or reported (actual) on the survey data file through the use of imputation codes. For more information, see the *Data File, Public Use: Public Libraries Survey: Fiscal Year 2001* (NCES 2003-398), forthcoming on the NCES web site. The following imputation rules were applied:

- A. For libraries that responded in 2000 but not 2001 (or in 1999 but not in 2000 or 2001):
 - A1. The growth rates were calculated for institutions that reported in both 2000 and 2001 (or in both 1999 and 2001). The mean (average) growth rate was calculated for each imputation cell.

- A2. The average changes computed in step A1 were applied to the 2000 data (or 1999 data) of 2001 nonresponding libraries to obtain an estimate for 2001.

This “growth rate” method was used for imputing centrals, branches, bookmobiles, ALA-MLS librarians, total librarians, total paid employees, book/serial volumes, subscriptions, audio, public service hours, library visits, reference transactions, total circulation, salaries, total staff expenditures, total collection expenditures, other operating expenditures, and total operating expenditures.

- A3. Employee benefits were derived by subtracting salaries from the estimated total staff expenditures determined in step A2.
- A4. Total operating expenditures were derived by summing total staff expenditures, total collection expenditures, and other operating expenditures estimated in step A2.
- A5. For income variables (i.e., total income and income from federal, state, and local government sources) both responding and nonresponding libraries in an imputation cell were arranged in decreasing order of size of population served. A nonresponding library’s 2000 (or 1999) data were pulled forward, and a growth rate was determined by calculating the growth of the next smallest library to the nonresponding library that had data for both 2001 and 2000 (or 2001 and 1999). This growth rate was applied to the nonresponding library’s 2000 (or 1999) data to obtain an estimate for 2001. If no prior year growth rate was available for the next smallest library, the growth rate was assumed to be 1.00.
- A6. Other income was derived by subtracting income from federal, state, and local sources from total income. Other paid employees was derived by subtracting librarians from total paid employees.
- A7. Children’s program attendance was estimated by multiplying the current-year total library visits by the prior-year ratio of children’s program attendance to total library visits.
- A8. Children’s circulation was estimated by multiplying the current-year total circulation by the prior-year ratio of children’s circulation to total circulation.
- A9. Materials in electronic format expenditures was set to zero if the sum of other operating expenditures and total collection expenditures equals zero.
- A10. For materials in electronic format and materials in electronic format expenditures, if only one of these variables needed imputation and if 2000 or 1999 reported data was greater than zero for both variables, the reported variable was multiplied by the ratio or inverse ratio to obtain an estimate for 2001. If the ratio/inverse ratio cannot be used but there was 2000 or 1999 reported data then that data was multiplied by the appropriate growth rate to obtain an estimate for 2001.
- A11. If the value for materials in electronic format was imputed to be zero and the value for materials in electronic format expenditures needs imputing, it was set to zero.
- A12. Access to electronic services and access to the Internet were imputed based on the prior-year response of “Yes” or “No”.
- A13. For electronic access expenditures, if access to electronic services was “No” and access to the Internet was “No”, or the sum of other operating expenditures plus total collection expenditures was zero, then electronic access expenditures’ imputed value was zero. Otherwise, the current year sum of other operating expenditures times the prior year ratio of electronic access expenditures to the prior year sum of other operating expenditures plus total collection expenditures was the imputed value. If only electronic access expenditures

was reported in the prior year (2000 or 1999), then it was multiplied by the appropriate growth rate to obtain an estimate for 2001.

- A14. If access to the Internet is "No", the imputed value of Internet use code is "NA" (not applicable). If the value for access to the Internet is "Yes", the imputed value for Internet use code is the prior-year value. If there is no prior year value, the imputed value is equal to the most frequent response in the same state.
- A15. For Internet terminals used by the general public and Internet terminals used by staff only, if there was 2000 or 1999 reported data, the data were carried forward as an estimate for 2001.
- A16. For population variables, the prior year data were carried forward in the current year.
- B. For libraries with no reported data in 1999, 2000, or 2001:
 - B1. The "growth rate" method (described in steps A1 and A2) was used to impute for 2001 if the prior year data (2000 or 1999) were imputed using prior year reported data and the imputed value was greater than zero.

This method was used for imputing centrals, branches, bookmobiles, ALA-MLS librarians, total librarians, total paid employees, book/serial volumes, subscriptions, audio, public service hours, reference transactions, circulation, total income, income from federal, state, and local sources, salaries, total staff expenditures, total collection expenditures, other operating expenditures, library materials in electronic format, operating expenditures for library materials in electronic format, and operating expenditures for electronic access.
 - B2. The mean of the imputation cell was calculated for all libraries that responded in 2001. The cell mean was adjusted for the size of a nonresponding library by multiplying it by the ratio of the nonrespondent's total population served to the mean size of population served for all responding libraries.

This method was used for imputing centrals, branches, bookmobiles, ALA-MLS librarians, total librarians, total paid employees, book/serial volumes, subscriptions, audio, public service hours, reference transactions, circulation, total income, income from federal, state, and local sources, salaries, total staff expenditures, total collection expenditures, other operating expenditures, library materials in electronic format, operating expenditures for library materials in electronic format, and operating expenditures for electronic access.
 - B3. To impute total library visits when the prior year imputed data were not based on reported data, library visits was summed over all responding libraries in an imputation cell, as was the population served. The ratio of total library visits to total population served was multiplied by the nonrespondent's population value to estimate the nonrespondent's library visits.
 - B4. Children's program attendance was estimated using the method described in step B3 where the ratio of total children's program attendance to total library visits for the responding libraries in an imputation cell was multiplied by the nonrespondent's current-year library visits.
 - B5. Children's circulation was estimated by calculating the ratio of children's circulation to total circulation for the responding libraries in an imputation cell and multiplying the ratio by the current-year total circulation of the nonresponding library.
 - B6. Employee benefits were derived by subtracting salaries from the estimated total staff expenditures determined in step B1 (or step B2). Other income was derived by subtracting income from federal, state, and local sources from total income. Other paid employees was derived by subtracting librarians from total paid employees.

- B7. Total operating expenditures were derived by summing total collection expenditures, total staff expenditures, and other operating expenditures estimated in step B1 (or step B2).
 - B8. Access to electronic services and access to the Internet were imputed based on the current-year response to operating expenditures for electronic access. (The value was set to "Yes" if electronic access expenditures was greater than 0; otherwise, the value was set to "No".)
 - B9. If there was no prior year value for Internet use, the imputed value is equal to the most frequent response for that variable in the same state.
 - B10. The median of the imputation cell was calculated for all libraries that responded in 2001. The cell median was not adjusted. This method was used for imputing Internet terminals used by the general public and Internet terminals used by staff only when there was no reported prior year (2000 or 1999) data. If the cell median was zero but based on the value of other electronic data items it was determined that the value should be greater than zero then the imputed value was equal to the unadjusted cell mean.
- C. For all nonresponding libraries:
- C1. Capital outlay was derived by imputing total expenditures (a derived variable which is the sum of total collections expenditures, total staff expenditures, other operating expenditures, and capital outlay) and subtracting total operating expenditures in order to get capital outlay. If the derived capital outlay had a negative value, it was changed to zero, total operating expenditures were changed to equal total expenditures, and total collection expenditures, total staff expenditures, and other operating expenditures were adjusted so that the sum would equal total operating expenditures. Alternatively, the cell mean (adjusted for population size) was used.
 - C2. The mean of the imputation cell was used to estimate videos and interlibrary loans. The cell mean was adjusted for the size of a nonresponding library by multiplying it by the ratio of the nonrespondent's total population served to the mean size of population served for all responding libraries.